

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-339101
 (43)Date of publication of application : 08.12.2000

(51)Int.Cl.

G06F 3/06
 G06F 12/08

(21)Application number : 11-149327
 (22)Date of filing : 28.05.1999

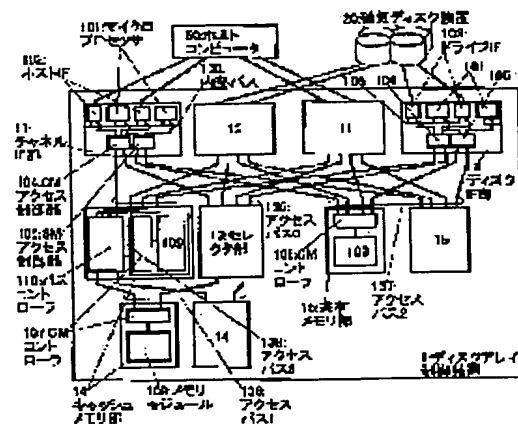
(71)Applicant : HITACHI LTD
 (72)Inventor : FUJIMOTO KAZUHIKA

(54) DISK ARRAY CONTROLLER

(57)Abstract:

PROBLEM TO BE SOLVED: To increase the capacity of a cache memory without increasing the number of PK and the number of wirings on a back plane and to improve data transfer throughput by installing cache memory parts on a package(PK) similar to a selector part and PK different from the selector part.

SOLUTION: New cache memory parts 14 (a CM controller 107 and a memory module 109) are installed on a selector PK. The capacity of a cache memory can be increased without increasing a memory PK. The number of cache memory parts 14 increases two times and the number of access paths between a path controller 110 and the CM controller 107 accordingly increases two times. Thus, data transfer throughput improves about two times. Even if the number of access paths increases two times, the increased access paths are required by the number of wirings on the selector PK and therefore it is not necessary to increase the number of wirings on a back plane.



LEGAL STATUS

[Date of request for examination]
 [Date of sending the examiner's decision of rejection]
 [Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
 [Date of final disposal for application]
 [Patent number]
 [Date of registration]
 [Number of appeal against examiner's decision of rejection]
 [Date of requesting appeal against examiner's decision of rejection]
 [Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

BEST AVAILABLE COPY

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-339101

(P2000-339101A)

(43) 公開日 平成12年12月8日 (2000.12.8)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード* (参考)
G 0 6 F 3/06	3 0 2	G 0 6 F 3/06	3 0 2 A 5 B 0 0 5
	5 4 0		5 4 0 5 B 0 6 5
12/08		12/08	G
			B
	3 2 0		3 2 0
審査請求 未請求 請求項の数 8 O L (全 13 頁)			

(21) 出願番号 特願平11-149327

(22) 出願日 平成11年5月28日 (1999.5.28)

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 藤本 和久

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(74) 代理人 100075096

弁理士 作田 康夫

Fターム(参考) 5B005 JJ11 KK12 MM05 MM11 NN12

5B065 BA01 CA30 CE11 CH01

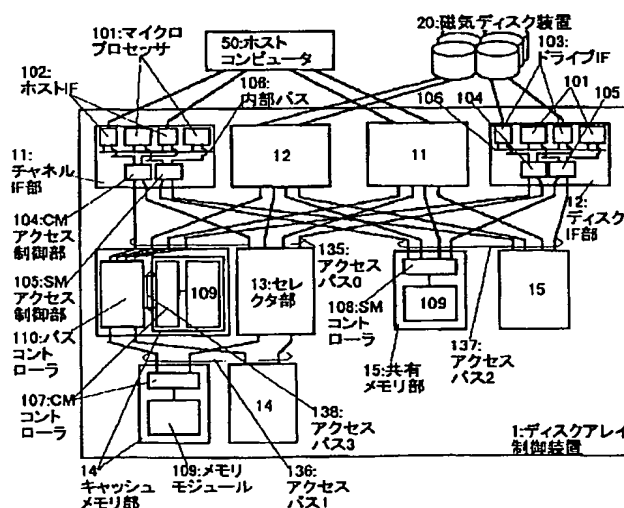
(54) 【発明の名称】 ディスクアレイ制御装置

(57) 【要約】

【課題】 パッケージ数およびバックプレーン上の配線数を増やすことなく、大容量キャッシュメモリを搭載でき、データ転送スループットの高いディスクアレイ制御装置を提供する。

【解決手段】 セレクタパッケージ上に新たなキャッシュメモリ部 (CMコントローラ107とメモリモジュール109) を設けることにより、メモリパッケージを増やすことなく、キャッシュメモリの容量を増やす。

図1



【請求項3】 ホストコンピュータとのインターフェースを有する複数のチャンネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッ

【請求項5】 ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、ディスクアレイ制御装置に関する制御情報を格納する共有メモリ部を有し、各チャネルインターフェース部は、前記ホストコンピュータと前記キャッシュメモリ部との間のデータ転送を実行し、各ディスク

インターフェース部は、前記磁気ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行し、前記複数のチャンネルインターフェース部および前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はスイッチを用いた相互結合網によって接続され、前記複数のチャンネルインターフェース部および前記複数のディスクインターフェース部と前記共有メモリ部との間は、それぞれ直接接続されたディスクアレイ制御装置において、前記キャッシュメモリ部が、前記複数のスイッチすべてに接続されるキャッシュメモリ部と、前記複数のスイッチの一部にのみ接続されるキャッシュメモリ部に分かれていることを特徴とするディスクアレイ制御装置。

【請求項6】ホストコンピュータとのインターフェースを有する複数のチャンネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記磁気ディスク装置に対しリード/ライトされるデータを一時的に格納するキャッシュメモリ部と、ディスクアレイ制御装置に関する制御情報を格納する共有メモリ部を有し、各チャンネルインターフェース部は、前記ホストコンピュータと前記キャッシュメモリ部との間のデータ転送を実行し、各ディスクインターフェース部は、前記磁気ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行し、前記複数のチャンネルインターフェース部および前記複数のディスクインターフェース部と前記共有メモリ部との間はスイッチを用いた相互結合網によって接続され、前記複数のチャンネルインターフェース部および前記複数のディスクインターフェース部と前記共有メモリ部との間は、それぞれ直接接続されたディスクアレイ制御装置において、前記キャッシュメモリ部が、前記複数のチャンネルインターフェース部およびディスクインターフェース部のすべてからアクセス可能なキャッシュメモリ部と、前記複数のチャンネルインターフェース部およびディスクインターフェース部の一部からのみアクセス可能なキャッシュメモリ部に分かれていることを特徴とするディスクアレイ制御装置。

【請求項7】請求項1から3のいずれかに記載のディスクアレイ制御装置であって、前記チャンネルインターフェース部および前記ディスクインターフェース部は、それぞれマイクロプロセッサと、前記キャッシュメモリ部へのアクセスを制御するキャッシュメモリアクセス制御部と、前記共有メモリ部へのアクセスを制御する共有メモリアクセス制御部を有し、前記キャッシュメモリアクセス制御部と前記キャッシュメモリ部間はセクタ部を介して接続されており、前記共有メモリ部へのアクセス制御部と前記共有メモリ部間は直接接続されていることを特徴とするディスクアレイ制御装置。

【請求項8】請求項4から6のいずれかに記載のディスクアレイ制御装置であって、前記チャンネルインターフェ

ース部および前記ディスクインターフェース部は、それぞれマイクロプロセッサと、前記キャッシュメモリ部へのアクセスを制御するキャッシュメモリアクセス制御部と、前記共有メモリ部へのアクセスを制御する共有メモリアクセス制御部を有し、前記キャッシュメモリアクセス制御部と前記キャッシュメモリ部間は複数のスイッチを用いた相互結合網により接続されており、前記共有メモリ部へのアクセス制御部と前記共有メモリ部間は直接接続されていることを特徴とするディスクアレイ制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、データを分割して複数の磁気ディスク装置に格納するディスクアレイ装置に関する。

【0002】

【従来の技術】半導体記憶装置を記憶媒体とするコンピュータの主記憶のI/O性能に比べて、磁気ディスクを記憶媒体とするディスクサブシステム（以下「サブシステム」という。）のI/O性能は3～4桁程度小さく、従来からこの差を縮めること、すなわちサブシステムのI/O性能を向上させる努力がなされている。サブシステムのI/O性能を向上させるための1つの方法として、複数の磁気ディスク装置でサブシステムを構成し、データを複数の磁気ディスク装置に格納する、いわゆるディスクアレイと呼ばれるシステムが知られている。

【0003】図2は従来のディスクアレイ制御装置の構成を、図3はそのディスクアレイ制御装置の実装構成を示す。ホストコンピュータ50とディスクアレイ制御装置2との間のデータ転送を実行する複数のチャンネルIF部11と、磁気ディスク装置20とディスクアレイ制御装置2間のデータ転送を実行する複数のディスクIF部12と、磁気ディスク装置20のデータを一時的に格納するキャッシュメモリ部14と、ディスクアレイ制御装置2に関する制御情報（例えば、チャンネルIF部11およびディスクIF部12とキャッシュメモリ部14との間のデータ転送制御に関する情報）を格納する共有メモリ部15とを備え、キャッシュメモリ部14はセクタ部13内のバスコントローラ110を介して、また共有メモリ部15はバスコントローラ110を介さず直接、全てのチャンネルIF部11およびディスクIF部12からアクセス可能な構成となっている。

【0004】図3に示すように、チャンネルIF部11はチャンネルIFパッケージ（以下パッケージを単にPKと略記する）81に、ディスクIF部12はディスクIFPK82に、セクタ部13はセクタPK83に、キャッシュメモリ部14および共有メモリ部15はメモリPK84に実装し、各パッケージは1枚のバックプレーン80上に実装する。したがって、各パッケージ間を接続するアクセスバス0、1、2（図2の符号135、1

36, 137) はバックプレーン 80 上に配線する。

【0005】このディスクアレイ制御装置 2 では、チャンネル I/F 部 11 およびディスク I/F 部 12 と共有メモリ部 15 との間は 1 対 1 に接続され、チャンネル I/F 部 11 およびディスク I/F 部 12 とキャッシュメモリ部 14 との間はバスコントローラ 110 を介して接続される。以下、チャンネル I/F 部 11 およびディスク I/F 部 12 と共有メモリ部 15 との間の接続形態をスター接続、チャンネル I/F 部 11 およびディスク I/F 部 12 とキャッシュメモリ部 14 との間の接続形態を階層スター接続と呼ぶ。

【0006】このように共有メモリ部 15 への接続形態とキャッシュメモリ部 14 への接続形態を異ならせるのは、ホストコンピュータへの応答時間を短くするために共有メモリ部へのアクセス時間をなるべく短く抑え、ホストコンピュータとの間のデータ転送スループットの向上に対応するためキャッシュメモリ部へのスループットをなるべく高くするという 2 つの要件を満足させるためである。チャンネル I/F 部 11 およびディスク I/F 部 12 と共有メモリ部 15 との間をバスコントローラ 110 を介さず直接接続することにより、アクセス時間を短縮できる。また、チャンネル I/F 部 11 およびディスク I/F 部 12 をバスコントローラ 110 を介してキャッシュメモリ部 14 に接続することにより、バスコントローラ 110 を介さず直接接続する場合に比べてキャッシュメモリ部 14 へのアクセスパス数を増やすことが可能になり、スループットを向上させることができる。

【0007】チャンネル I/F 部 11 は、ホストコンピュータ 50 と接続するためのホスト I/F 102、ホストコンピュータ 50 に対する入出力を制御するマイクロプロセッサ 101、キャッシュメモリ部 14 へのアクセスを制御する CM アクセス制御部 104、および共有メモリ部 15 へのアクセスを制御する SM アクセス制御部 105 を有している。また、ディスク I/F 部 12 は、磁気ディスク装置 20 と接続するためのドライブ I/F 103、磁気ディスク装置 20 に対する入出力を制御するマイクロプロセッサ 101、キャッシュメモリ部 14 へのアクセスを制御する CM アクセス制御部 104、および共有メモリ部 15 へのアクセスを制御する SM アクセス制御部 105 を有している。また、ディスク I/F 部 12 は、RAID 機能の実行も行う。

【0008】

【発明が解決しようとする課題】近年、磁気ディスクの大容量化は目覚しく、このためディスクアレイの記憶容量も急激に増加している。それに伴って、ディスクアレイ制御装置に搭載するキャッシュメモリの容量を増加させる必要がある。また、ホストコンピュータとのインターフェースとして 100 MB/s の高スループットを有するファイバチャネルが採用されてきており、さらに今後数年でそのスループットは 2~4 倍に向上する。このような高スループットインターフェースに対応するた

め、ディスクアレイ制御装置のデータ転送スループットも大幅に増加させる必要がある。

【0009】図 2 に示すチャンネル I/F 部 11 およびディスク I/F 部 12 とキャッシュメモリ部 14 との間のスループット向上を目指した階層スター接続方式において、キャッシュメモリの容量を増やし、データ転送スループットを向上させるためには、キャッシュメモリ部 14 を実装するメモリ PK 84 (図 3) の数を増やす必要がある。しかしながら、図 2 に示したディスクアレイ制御装置 2 では、メモリ PK 84 の数を増やすことは筐体サイズを大きくすることになるため難しい。また、仮にメモリ PK 84 の数を増やせたとしても、セレクト部 13 とキャッシュメモリ部 14 の間のアクセスパス数が増え、各パッケージを実装するバックプレーン 80 上の各パッケージ間を結ぶ配線数が膨大になるという問題が生じる。

【0010】そこで本発明の目的は、階層スター接続方式を用いた従来のディスクアレイ制御装置において、パッケージ数やバックプレーン上の配線数を増やすことなく、キャッシュメモリ容量を増やし、データ転送スループットを向上させたディスクアレイ制御装置を提供することにある。

【0011】

【課題を解決するための手段】上記目的は、ホストコンピュータとのインターフェースを有する複数のチャンネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記磁気ディスク装置に対しリード/ライトされるデータを一時的に格納するキャッシュメモリ部と、ディスクアレイ制御装置に関する制御情報を格納する共有メモリ部を有し、各チャンネルインターフェース部は、前記ホストコンピュータと前記キャッシュメモリ部との間のデータ転送を実行し、各ディスクインターフェース部は、前記磁気ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行し、前記複数のチャンネルインターフェース部および前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセレクト部を介して接続され、前記複数のチャンネルインターフェース部および前記複数のディスクインターフェース部と前記共有メモリ部との間は、それぞれセレクト部を介さずに直接接続されたディスクアレイ制御装置において、前記キャッシュメモリ部が、前記セレクト部と同じパッケージ上に実装されるキャッシュメモリ部と、前記セレクト部と異なるパッケージ上に実装されるキャッシュメモリ部に分かれていることを特徴とするディスクアレイ制御装置によって達成される。

【0012】また上記目的は、ホストコンピュータとのインターフェースを有する複数のチャンネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記磁気デ

ィスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、ディスクアレイ制御装置に関する制御情報を格納する共有メモリ部を有し、各チャンネルインターフェース部は、前記ホストコンピュータと前記キャッシュメモリ部との間のデータ転送を実行し、各ディスクインターフェース部は、前記磁気ディスク装置と前記キャッシュメモリ部との間のデータ転送を実行し、前記複数のチャンネルインターフェース部および前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はスイッチを用いた相互結合網によって接続され、前記複数のチャンネルインターフェース部および前記複数のディスクインターフェース部と前記共有メモリ部との間は、それぞれ直接接続されたディスクアレイ制御装置において、前記キャッシュメモリ部が、前記相互結合網と同じパッケージ上に実装されるキャッシュメモリ部と、前記相互結合網と異なるパッケージ上に実装されるキャッシュメモリ部に分かれていることを特徴とするディスクアレイ制御装置によって達成される。

【0013】

【発明の実施の形態】（実施例1）図1、図3に本発明の一実施例を示す。ディスクアレイ制御装置1は、ホストコンピュータ50との2つのインターフェース部（チャンネルIF部）11と、磁気ディスク装置20との2つのインターフェース部（ディスクIF部）12と、2つのセクタ部13と、セクタ部13内のキャッシュメモリ部14と、セクタ部13の外の2つのキャッシュメモリ部14と、2つの共有メモリ部15と、アクセスバス0：135と、アクセスバス1：136と、アクセスバス2：137と、アクセスバス3：138を有する。

【0014】チャンネルIF部11は、ホストコンピュータ50との2つのIF（ホストIF）102と、2つのマイクロプロセッサ101と、キャッシュメモリ部14への1つのアクセス制御部（CMアクセス制御部）104と、共有メモリ部15への1つのアクセス制御部（SMアクセス制御部）105を有し、ホストコンピュータ50とキャッシュメモリ部14間のデータ転送、およびマイクロプロセッサ101と共有メモリ部15間の制御情報の転送を実行する。マイクロプロセッサ101およびホストIF102は内部バス106によって接続され、CMアクセス制御部104は2つのホストIF102に直接接続されている。また、SMアクセス制御部105は2つのマイクロプロセッサ101に直接接続されている。チャンネルIF部11の構成要素は図3のように、チャンネルIFPK81上に実装する。

【0015】ディスクIF部12は、磁気ディスク装置20との2つのIF（ドライブIF）103と、2つのマイクロプロセッサ101と、キャッシュメモリ部14への1つのアクセス制御部（CMアクセス制御部）10

4と、共有メモリ部15への1つのアクセス制御部（SMアクセス制御部）105を有し、磁気ディスク装置20とキャッシュメモリ部14間のデータ転送、およびマイクロプロセッサ101と共有メモリ部15間の制御情報の転送を実行する。

【0016】マイクロプロセッサ101およびドライブIF103は内部バス106によって接続され、CMアクセス制御部104は2つのドライブIF103に直接接続されている。また、SMアクセス制御部105は2つのマイクロプロセッサ101に直接接続されている。ディスクIF部12はRAID機能の実行も行う。ディスクIF部12の構成要素は図3のようにディスクIFPK82上に実装する。

【0017】キャッシュメモリ部14は、キャッシュメモリ（CM）コントローラ107とメモリモジュール109を有し、磁気ディスク装置20へ記録するデータを格納する。セクタ部13の外のキャッシュメモリ部14の構成要素はメモリPK84上に実装する。

【0018】共有メモリ部15は、共有メモリ（SM）コントローラ108とメモリモジュール109を有し、ディスクアレイ制御装置1の制御情報（例えば、チャンネルIF部およびディスクIF部とキャッシュメモリ部14との間のデータ転送制御に関する情報）等を格納する。共有メモリ部15の構成要素は図3のように、メモリPK84上に実装する。

【0019】セクタ部13は、バスコントローラ110とキャッシュメモリ部14（CMコントローラ107とメモリモジュール109）を有し、CMコントローラ107へのアクセスの振り分け、および磁気ディスク装置20へ記録するデータの格納を行う。セクタ部13の構成要素はセクタPK83上に実装する。

【0020】上記各パッケージは、図3に示すように1つのバックプレーン80上に実装する。

【0021】CMアクセス制御部104には2本のアクセスバス0：135を接続し、それらを2つの異なるセクタ部13内のバスコントローラ110にそれぞれ接続する。バスコントローラ110には2本のアクセスバス1：136を接続し、それらを2つの異なるメモリPK84上のキャッシュメモリ部14内のCMコントローラ107にそれぞれ接続する。また、セクタPK83上のキャッシュメモリ部14内のCMコントローラ107への2本のアクセスバス3：138を接続する。したがってメモリPK84上のキャッシュメモリ部14内のCMコントローラ107には、2つのセクタ部13から1本ずつ、計2本のアクセスバス1：136が接続される。こうすることにより、CMアクセス制御部104からメモリPK84上のキャッシュメモリ部14内のCMコントローラ107へは、2つのセクタ部13のどちらを経由してもアクセスが可能となる。これにより、1つのアクセスバスまたはバスコントローラ110に障

害が発生した場合でも、もう1つのアクセスルートによりキャッシュメモリ部14へアクセスすることが可能となるため、耐障害性を向上させることができる。

【0022】SMアクセス制御部105には2本のアクセスバス2:137を接続し、そのアクセスバス2:137を2つの異なるSMコントローラにそれぞれ接続する。したがってSMコントローラ108には、2つのチャンネルIF部11および2つのディスクIF部12から1本ずつ、計4本のアクセスバス2:137が接続される。

【0023】本実施例では、1つのSMアクセス制御部105と1つのSMコントローラ108の間には1本のアクセスバス2:137を接続したが、このアクセスバス2:137を2本に増やすことにより、1つのSMアクセス制御部105から1つのSMコントローラ108へのアクセスルートが2つとなるため、耐障害性を向上させることができる。ここで、上記個数は一実施例にすぎず、個数を上記に限定するものではない。

【0024】バスコントローラ110には、2つのチャンネルIF部11と、2つのディスクIF部12からそれぞれ1本ずつ、計4本のアクセスバス0:135が接続される。また、バスコントローラ110には、メモリPK84上のキャッシュメモリ部14へのアクセスバス1:136が1本ずつと、セクタPK83上のキャッシュメモリ部14への2本のアクセスバス3:138の、計4本が接続される。

【0025】本発明の大きな特徴は、セクタPK83上に新たなキャッシュメモリ部14（CMコントローラ107とメモリモジュール109）を設ける点にある。こうすることにより、メモリPK84を増やすことなく、キャッシュメモリの容量を増やすことができる。また、キャッシュメモリ部14の数が2倍に増え、それに伴いバスコントローラ110とCMコントローラ107との間のアクセスバス数も2倍に増える。したがって、データ転送スループットが約2倍に向上する。バスコントローラ110とCMコントローラ107との間のアクセスバス数が2倍に増えても、増加分のアクセスバスはセクタPK83上の配線だけですむため、バックプレーン80上の配線数を増やす必要がない。したがって、図2に示す従来のディスクアレイ制御装置2において、セクタ部13内の構成を変えるだけでキャッシュメモリ容量およびデータ転送スループットを増やすことが可能となる。

【0026】図4は、CMアクセス制御部104内の構成を示している。CMアクセス制御部104は、セクタ302と、アドレス、コマンド、データを一時格納するバッファ303と、セクタ部13に繋がるアクセスバス0:135とのバスIF301と、データのエラーチェック部300と、データ転送制御部310を有する。

【0027】セクタ302の2つのポートはデータ線210でホストIF102あるいはドライブIF103に接続される。また、セクタ302の他の2つのポートはバスIF301に接続される。バスIF301はアクセスバス0:135でセクタ部13内のバスコントローラ110に接続される。

【0028】データ転送制御部310は、制御線1:211でホストIF102あるいはドライブIF103に接続され、制御線2:212でバスコントローラ110内のデータ転送制御部315（図6で後述）に接続される。また、データ転送制御部310は、アービタ308によりホストIF102あるいはドライブIF103からのアクセス要求のアービトレーションを行い、セクタ302の切り替えを行う。

【0029】図6は、セクタ部13内のバスコントローラ110の構成を示している。バスコントローラ110は、チャンネルIF部11およびディスクIF部12に繋がるアクセスバス0:135との4つのバスIF301と、メモリPK84上のキャッシュメモリ部14に繋がるアクセスバス1:136との2つのバスIF301およびセクタPK83上のキャッシュメモリ部14に繋がるアクセスバス3:138との2つのバスIF301と、両者間を互いに接続するセクタ306と、パケットバッファ303と、データのエラーチェック部300と、CMアクセス制御部104から送出されたアドレスおよびコマンドを解析するアドレス・コマンド（addr-cmd）解析部305と、データ転送制御部315を有する。

【0030】データ転送制御部315は、制御線2:212でCMアクセス制御部104内のデータ転送制御部310に接続され、制御線3:213でCMコントローラ107内のデータ転送制御部315に接続される。また、データ転送制御部315は、アービタ308により、addr-cmd解析部305で解析した4本のアクセスバス0:135からのアクセス要求のアービトレーションを行い、セクタ306の切り替えを行う。

【0031】バッファ303は、アクセスバス0:135側のバスとアクセスバス1:136およびアクセスバス3:138側のバスでデータ転送速度に差がある場合、速度差を吸収するために、転送するデータの一部または全部をバッファリングする。

【0032】addr-cmd解析部305は、アドレスおよびコマンドを格納するバッファと、アドレス抽出部と、コマンド抽出部を有する（図示略）。このaddr-cmd解析部305では、CMアクセス制御部104に接続される4本のアクセスバス0:135それぞれに1つずつ割り当てられたバッファに、アドレス、コマンドを格納する。アドレス抽出部およびコマンド抽出部では、アクセスするCMコントローラ107とアクセスの種類を割り出し、データ転送制御部315内のアービタ

308へ送出する。

【0033】図7は、キャッシュメモリ部14内の構成を示している。CMコントローラ107は、バスコントローラ110に繋がるアクセスバス1:136またはアクセスバス3:138との2つのバスIF301と、セクタ304と、データを一時格納するバッファ303と、データのエラーチェック部300と、メモリモジュール109へのアクセスを制御するメモリ制御部307と、CMアクセス制御部104から送出されたアドレスおよびコマンドを解析するaddr-cmd解析部305と、データ転送制御部315を有する。

【0034】データ転送制御部315は、制御線3:213でバスコントローラ110内のデータ転送制御部315に接続され、制御線4:214でメモリ制御部107に接続される。また、データ転送制御部315は、アービタ308により、addr-cmd解析部305で解析した2本のアクセスバスからのアクセス要求のアービトレーションを行い、セクタ304の切り替えを行う。

【0035】addr-cmd解析部305は、アドレスおよびコマンドを格納するバッファおよびアドレス抽出部、コマンド抽出部を有する(図示略)。addr-cmd解析部305では、CMコントローラ107に接続される2本のアクセスバスそれぞれに1つずつ割り当てられたバッファに、アドレス、コマンドを格納する。アドレス抽出部およびコマンド抽出部では、アクセスするメモリのアドレスとアクセスの種類を割り出し、メモリ制御部307へ送出する。また、2本のアクセスバスからのアクセス要求をデータ転送制御部315内のアービタ308へ送出する。

【0036】次に、キャッシュメモリ部14へのアクセス時の手順について述べる。キャッシュメモリ部14へアクセスする場合、マイクロプロセッサ101は、ホストIF102あるいはドライブIF103へ、キャッシュメモリ部14へのアクセス開始を指示する。

【0037】アクセス開始の指示を受けたホストIF102あるいはドライブIF103は、制御線1:211によりCMアクセス制御部104内のデータ転送制御部310へアクセス開始を示す信号を送出する。それとともに、データ線210を通してアドレス、コマンド、データ(データの書き込み時のみ)を送出する。

【0038】CMアクセス制御部104は、データ線210を通して送られてきたアドレス、コマンド、データ(データの書き込み時のみ)をバッファ303に格納する。データ転送制御部310はアービトレーションを行ってバスIF301の使用権を決定し、セクタ302を切り替える。

【0039】図9は、キャッシュメモリ部14へデータを書き込む場合の、CMアクセス制御部104からCMコントローラ107へのアクセスの流れを示している。

CMアクセス制御部104内のデータ転送制御部310は、アービトレーションによってアクセスバス0:135の使用権が決定されると、制御線2:212によってバスコントローラ110内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ501)。続いて、アドレスおよびコマンドを送出する(ステップ502)。

【0040】バスコントローラ110内のデータ転送制御部315は、CMアクセス制御部104からREQ信号を受け取ると、次にアクセスバス0:135を通して送られてくるアドレスおよびコマンドを受信し、addr-cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行う(ステップ503)。アービトレーションの結果、アクセスバス1:136またはアクセスバス3:138への接続権を得たら、データ転送制御部315はセクタ306を切り替える(ステップ504)とともに、制御線2:212により、CMアクセス制御部104内のデータ転送制御部310へ、アクセスバス1:136またはアクセスバス3:138への接続権が得られたことを示す信号(ACK)を返す(ステップ505)。

【0041】次にデータ転送制御部315は、制御線3:213によってCMコントローラ107内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ506)。続いて、アドレスおよびコマンドを送出する(ステップ507)。

【0042】CMアクセス制御部104はACK信号を受けると、バッファ303からデータを読み出し、セクタ302、バスIF301を介してアクセスバス0:135へ送出する。バスコントローラ110は、アクセスバス0:135を通して送られてきたデータを、バスIF301およびセクタ306を介してアクセスバス1:136またはアクセスバス3:138へ送出する(ステップ509)。

【0043】CMコントローラ107内のデータ転送制御部315は、制御線3:213によってREQ信号を受け取ると、次にアクセスバス1:136またはアクセスバス3:138を通して送られてくるアドレスおよびコマンドを受信し、addr-cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行い(ステップ508)、セクタ304を切り替える。アクセスバス1:136またはアクセスバス3:138を通して送られてくるデータはバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアクセスのための前処理を行う(ステップ510)。次に、バッファ303からデータを読み出し、セクタ304を介してメモリモジュール109へ書き込む(ステップ511)。

【0044】メモリモジュール109へのアクセスが終了すると、メモリアクセスの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス

(STATUS)を生成する(ステップ512)。次に、ステータスをバスコントローラ110を介してCMアクセス制御部104へ送出する(ステップ513)。バスコントローラ110内のデータ転送制御部315はステータスを受け取ると、CMコントローラ107へのREQ信号をオフする(ステップ514)。

【0045】CMアクセス制御部104内のデータ転送制御部310はステータスを受け取ると、バスコントローラ110へのREQ信号をオフする(ステップ515)。バスコントローラ110内のデータ転送制御部315はCMアクセス制御部104からのREQ信号のオフを確認すると、CMアクセス制御部104へのACK信号をオフする(ステップ516)。

【0046】CMアクセス制御部104内のデータ転送制御部310はステータスを受け取ると、制御線1:211により、ホストIF102あるいはドライブIF103に、キャッシュメモリ部14へのアクセスの終了を報告する。

【0047】キャッシュメモリ部14からデータを読み出す場合の、CMアクセス制御部104からCMコントローラ107へのアクセスの流れは、ステップ501から508までとステップ512以降は、データの書き込みの場合と同じである。

【0048】CMアクセス制御部104はステップ505でACK信号を受けると、データの受信待ち状態に入る。ステップ508でメモリアクセス権を得ると、CMコントローラ107はメモリモジュール109からデータを読み出し、セクタ304、バスIF301を介してアクセスバス1:136またはアクセスバス3:138にデータを送出する。バスコントローラ110は、アクセスバス1:136またはアクセスバス3:138を通してデータを受信すると、バスIF301およびセクタ306を介してアクセスバス0:135にデータを送出する。CMアクセス制御部104は、アクセスバス0:135を通してデータを受信すると、セクタ302、データ線210を介してホストIF102あるいはドライブIF103へデータを送出する。

【0049】図5は、SMアクセス制御部105内の構成を示している。SMアクセス制御部104は、セクタ302と、アドレス、コマンド、データを一時格納するバケットバッファ303と、SMコントローラ108に繋がるアクセスバス2:137とのバスIF301と、データのエラーチェック部300と、データ転送制御部310を有する。

【0050】セクタ302の2つのポートはデータ線220で2つのマイクロプロセッサ101にそれぞれ接続される。また、セクタ302の他の2つのポートは

2つのバスIF301にそれぞれ接続される。バスIF301はアクセスバス2:137でSMコントローラ108に接続される。データ転送制御部310は、制御線5:221でマイクロプロセッサ101に接続され、制御線6:222でSMコントローラ108内のデータ転送制御部315に接続される。また、データ転送制御部310は、アービタ308によりマイクロプロセッサ101からのアクセス要求のアービトレーションを行い、セクタ302の切り替えを行う。

【0051】図8は、共有メモリ部15内の構成を示している。共有メモリ部15は、SMコントローラ108とメモリモジュール109を有する。SMコントローラ108は、SMアクセス制御部105に繋がるアクセスバス2:137との4つのバスIF301と、セクタ309と、データを一時格納するバケットバッファ303と、データのエラーチェック部300と、メモリモジュール109へのアクセスを制御するメモリ制御部307と、SMアクセス制御部105から送出されたアドレスおよびコマンドを解析するadr-cmd解析部305と、データ転送制御部315を有する。

【0052】データ転送制御部315は、制御線6:222でSMアクセス制御部105内のデータ転送制御部310に接続され、制御線4:214でメモリ制御部307に接続される。また、データ転送制御部315は、アービタ308により、adr-cmd解析部305で解析した4本のアクセスバス2:137からのアクセス要求のアービトレーションを行い、セクタ309の切り替えを行う。

【0053】adr-cmd解析部305は、アドレスおよびコマンドを格納するバッファと、アドレス抽出部と、コマンド抽出部を有する(図示略)。adr-cmd解析部305では、SMコントローラ108に接続される4本のアクセスバス2:137それぞれに1つずつ割り当てられたバッファに、アドレス、コマンドを格納する。アドレス抽出部およびコマンド抽出部では、アクセスするメモリのアドレスとアクセスの種類を割り出し、メモリ制御部307へ送出する。また、4本のアクセスバス2:137からのアクセス要求をデータ転送制御部315内のアービタ308へ送出する。

【0054】次に、共有メモリ部15へのアクセス時の手順について述べる。共有メモリ部15へアクセスする場合、マイクロプロセッサ101は、制御線5:221によりSMアクセス制御部105内のデータ転送制御部310へアクセス開始を示す信号を送出する。それとともに、データ線220を通してアドレス、コマンド、データ(データの書き込み時のみ)を送出する。

【0055】SMアクセス制御部105は、データ線220を通して送られてきたアドレス、コマンド、データ(データの書き込み時のみ)をバケットバッファ303に格納する。データ転送制御部310はアービトレーシ

ョンを行ってバスIF301の使用権を決定し、セクタ302を切り替える。

【0056】図10は、共有メモリ部15へデータを書き込む場合の、SMアクセス制御部105からSMコントローラ108へのアクセスの流れを示している。SMアクセス制御部105内のデータ転送制御部310は、アービトレーションによってアクセスバス2:137の使用権が決定されると、制御線6:222によってSMコントローラ108へアクセス開始を示す信号(REQ)を出す(ステップ601)。続いて、アドレス、コマンド、およびデータを連続して送出する(ステップ602)。

【0057】SMコントローラ108内のデータ転送制御部315は、制御線6:222によってREQ信号を受け取ると、次にアクセスバス2:137を通して送られてくるアドレス、コマンド、およびデータを受信する。アドレスとコマンドは、adr-cmd解析部305で解析し、アクセス要求に基づいてアービトレーションを行い(ステップ603)、セクタ309を切り替える。データはパケットバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアクセスのための前処理を行う(ステップ604)。次に、パケットバッファ303からデータを読み出し、セクタ309を介してメモリモジュール109へ書き込む(ステップ605)。

【0058】メモリモジュール109へのアクセスが終了すると、メモリアクセスの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス(STATUS)を生成する(ステップ606)。次に、ステータスをSMアクセス制御部105へ送出する(ステップ607)。SMアクセス制御部105内のデータ転送制御部310はステータスを受け取ると、SMコントローラ108へのREQ信号をオフする(ステップ608)。

【0059】SMアクセス制御部105内のデータ転送制御部310はステータスを受け取ると、制御線5:221により、マイクロプロセッサ101へ共有メモリ部15へのアクセスの終了を報告する。

【0060】共有メモリ部15からデータを読み出す場合のSMアクセス制御部105からSMコントローラ108へのアクセスの流れは、ステップ601から604までとステップ606以降は、データの書き込みの場合と同じである。

【0061】ステップ604でメモリアクセスの前処理を行った後、SMコントローラ108はメモリモジュール109からデータを読み出し、セクタ309、バスIF301を介してアクセスバス2:137にデータを送出する。

【0062】SMアクセス制御部105は、アクセスバ

ス2:137を通してデータを受信すると、セクタ302、データ線220を介してマイクロプロセッサ101へデータを送出する。

【0063】本実施例によれば、図2に示す従来のディスクアレイ制御装置2において、セクタ部13の構成を変える(セクタPK83内にキャッシュメモリ部14を設ける)だけでキャッシュメモリ容量およびデータ転送スループットを増やすことが可能となる。

【0064】(実施例2)図12に本発明の他の実施例を示す。本実施例ではチャンネルIF部11、ディスクIF部12、セクタ部13、およびセクタ部13の外のキャッシュメモリ部の数を図1のディスクアレイ制御装置の倍に増やしている。チャンネルIF部11、ディスクIF部12、セクタ部13、キャッシュメモリ部14、および共有メモリ部15内の構成は図示していないが、図1に示すディスクアレイ制御装置と同様である。

【0065】チャンネルIF部11は例えば図3のチャンネルIFPK81上に、ディスクIF部12はディスクIFPK82上に、セクタ部13はセクタPK83上に、セクタ部13の外のキャッシュメモリ部14および共有メモリ部15はメモリPK84上に実装し、各パッケージ(PK)をバックプレーン80に実装する。各PK間を接続するアクセスバスはバックプレーン80上に配線する。

【0066】データを転送するアクセスバス0:135、アクセスバス1:136のビット幅は制御情報を転送するアクセスバス2:137のビット幅に比べて2~4倍広い。したがって、ディスクアレイ制御装置1に搭載するチャンネルIF部11およびディスクIF部12の数を増やす場合、セクタ部13内のバスコントローラ110のLSIのピン数ネックを避けるためセクタ部13の数も増やす必要がある。

【0067】そうした場合、すべてのチャンネルIF部11およびディスクIF部12をすべてのセクタ部13に接続することができない。したがって本実施例では、図12の上方の2つのチャンネルIF部11と2つのディスクIF部12はその下の2つのセクタ部13に接続する。また、図12の下方の2つのチャンネルIF部11と2つのディスクIF部12はその上の2つのセクタ部13に接続する。

【0068】上記のような接続になるため、セクタ部13内のキャッシュメモリ部14にアクセスできるチャンネルIF部11およびディスクIF部12は限定される。本実施例では、図の上方の各IF部はその下のセクタ部13内のキャッシュメモリ部14にアクセス可能で、図の下方の各IF部はその上のセクタ部13内のキャッシュメモリ部14にアクセス可能である。

【0069】ディスクアレイ制御装置1では、耐障害性の向上の観点から、あるチャンネルIF部11またはディスクIF部12からあるキャッシュメモリ部14にアク

セスできなくなった場合、他のチャンネル I F 部 11 またはディスク I F 部 12 からそのキャッシュメモリ部 14 にアクセスできる構成としなければならない。セレクト部 13 の外のキャッシュメモリ部 14 はすべての I F 部からアクセス可能であるので問題ない。

【0070】しかしながら、セレクト部 13 内のキャッシュメモリ部 14 へはアクセスできる I F 部が限られるので、以下のような対策を施す。すなわち図 12 に示すように、ディスクアレイ制御装置 1 内を電源境界 800 で 2 つに分ける。電源境界 800 を挟んだ 2 つの領域にはそれぞれ異なる電源から電力を供給する。セレクト部 13 内のキャッシュメモリ部 14 を電源境界 800 を挟んだ 2 つの領域内のものどうして二重化する。ここで図の上方の 2 つのセレクト部 13 内のキャッシュメモリ部 14 とうしと図の下方の 2 つのセレクト部 13 内のキャッシュメモリ部 14 とうしはそれぞれ別個に二重化する。さらに、チャンネル I F 部 11 またはディスク I F 部 12 からセレクト部 13 への 2 本のアクセスバス 0:135 を電源境界 800 を挟んだ 2 つのセレクト部 13 にそれぞれ 1 本ずつ接続する。

【0071】上記のようにすることで、アクセスバス 0:135 の障害で 1 つのチャンネル I F 部 11 またはディスク I F 部 12 から 1 つのセレクト部 13 内のキャッシュメモリ部 14 にアクセスできなくなった場合に、同じセレクト部 13 に繋がる別のチャンネル I F 部 11 またはディスク I F 部 12 から該キャッシュメモリ部 14 にアクセス可能となる。また、電源境界 800 を挟んだ 2 つの領域の一方の供給電源が障害等で遮断された場合、供給電源の異なる領域のセレクト部 13 内のキャッシュメモリ部 14 で装置の運転を継続することが可能となる。

【0072】セレクト部 13 の外のキャッシュメモリ部 14 および共有メモリ部 15 も電源境界 800 を挟んだ 2 つの領域内のものどうして二重化する。このようにすることで、電源遮断時に上記と同じ効果が得られる。

【0073】（実施例 3）図 11 に本発明の他の実施例を示す。本実施例では、チャンネル I F 部 11 およびディスク I F 部 12 とキャッシュメモリ部 14 間をスイッチ (SW) 16 を用いた相互結合網 140 で接続する。この場合にも、図 1 に示したバスコントローラ 110 を介して接続した構成と同様に、セレクト部 13 上に相互結合網 140 とキャッシュメモリ部 14 (CM コントローラ 107 とメモリモジュール 109) を有するスイッチ部 17 を実装し、CM コントローラ 107 を相互結合網 140 に接続することで、新たなメモリ PK 84 を増やすことなく、またバックプレーン 80 上の配線数を増やすことなく、キャッシュメモリ容量およびデータ転送

スループットを増やすことが可能となる。

【0074】

【発明の効果】本発明によれば、新たなキャッシュメモリ用パッケージを増やすことなく、またバックプレーン上の配線数を増やすことなく、従来のディスクアレイ制御装置のキャッシュメモリ容量およびデータ転送スループットを増やすことが可能となる。これによって、大容量キャッシュメモリを搭載し、データ転送スループットの高いディスクアレイ制御装置を提供することが可能となる。

【図面の簡単な説明】

【図 1】本発明の一実施例のディスクアレイ制御装置の構成を示すブロック図。

【図 2】従来例のディスクアレイ制御装置の構成を示すブロック図。

【図 3】従来例のディスクアレイ制御装置の実装構成を示すブロック図。

【図 4】本発明の一実施例の CM アクセス制御部の構成を示すブロック図。

【図 5】本発明の一実施例の SM アクセス制御部の構成を示すブロック図。

【図 6】本発明の一実施例のバスコントローラの構成を示すブロック図。

【図 7】本発明の一実施例のキャッシュメモリ部の構成を示すブロック図。

【図 8】本発明の一実施例の共有メモリ部の構成を示すブロック図。

【図 9】キャッシュメモリ部へのデータの書き込み時の手順を示すシーケンス図。

【図 10】共有メモリ部へのデータの書き込み時の手順を示すシーケンス図。

【図 11】本発明のディスクアレイ制御装置の他の構成例を示すブロック図。

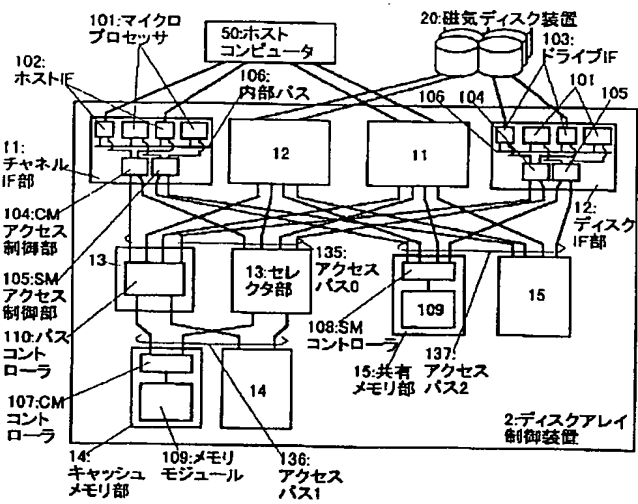
【図 12】本発明のディスクアレイ制御装置の他の構成例を示すブロック図。

【符号の説明】

1…ディスクアレイ制御装置、11…チャンネル I F 部、12…ディスク I F 部、13…セレクト部、14…キャッシュメモリ部、15…共有メモリ部、20…磁気ディスク装置、50…ホストコンピュータ、101…マイクロプロセッサ、102…ホスト I F、103…ドライブ I F、104…CM アクセス制御部、105…SM アクセス制御部、106…内部バス、107…CM コントローラ、108…SM コントローラ、109…メモリモジュール、110…バスコントローラ、135…アクセスバス 0、136…アクセスバス 1、137…アクセスバス 2、138…アクセスバス 3。

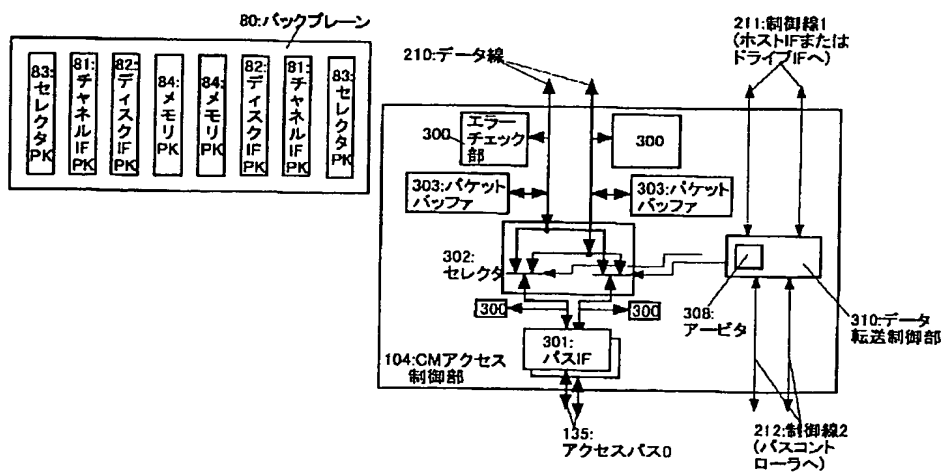
【図 2】

2



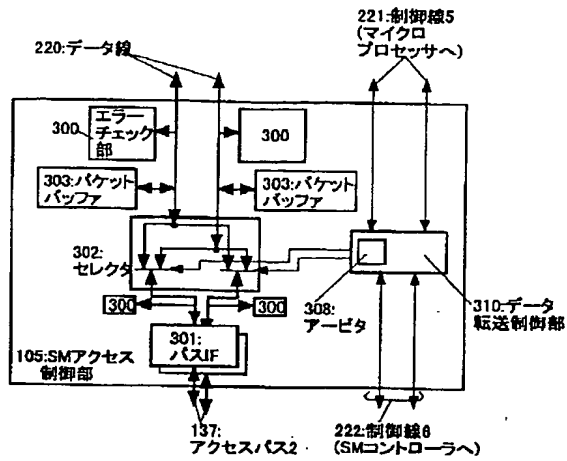
【図4】

图4



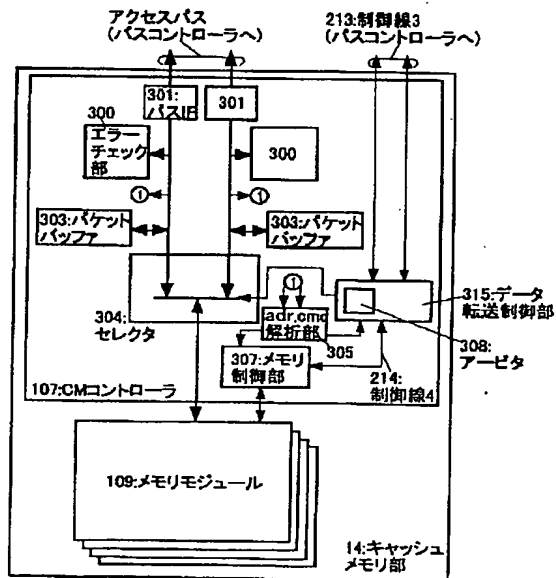
【図5】

図5



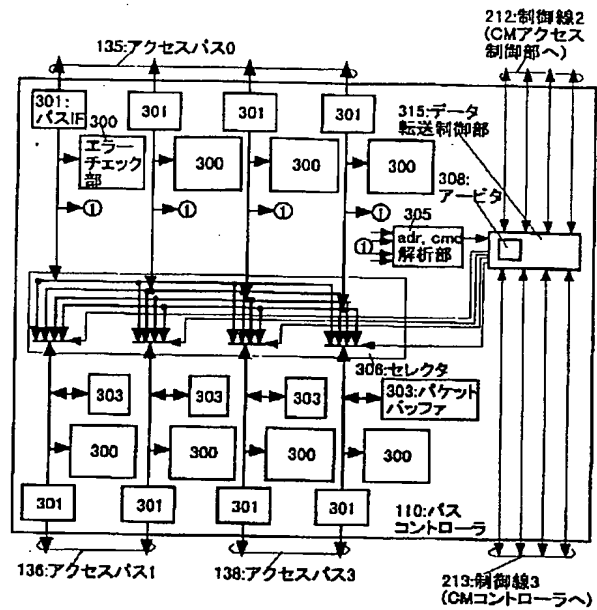
【図7】

図7



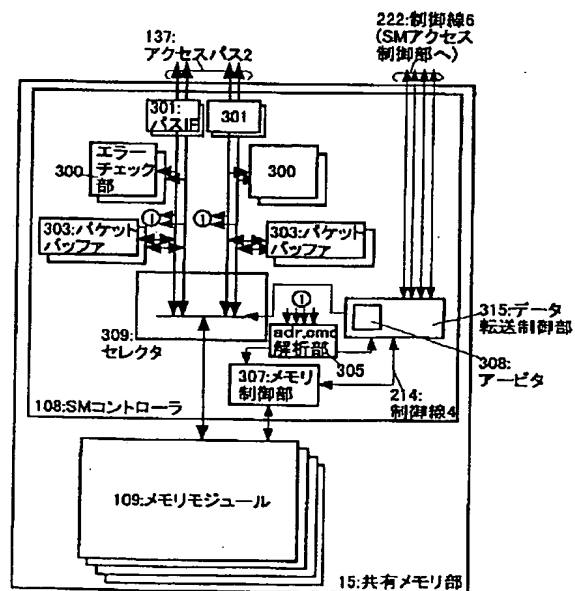
【図6】

図6



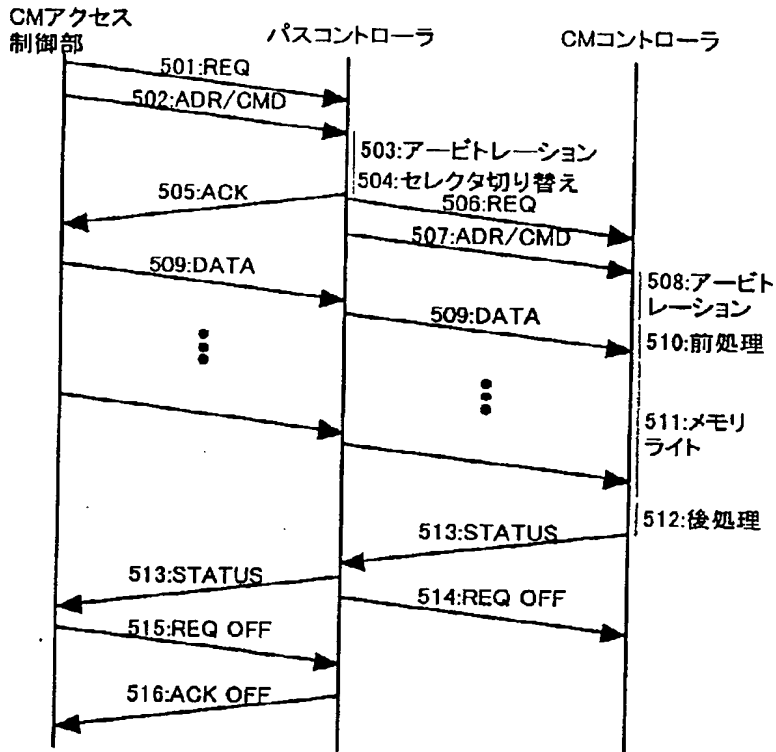
【図8】

図8



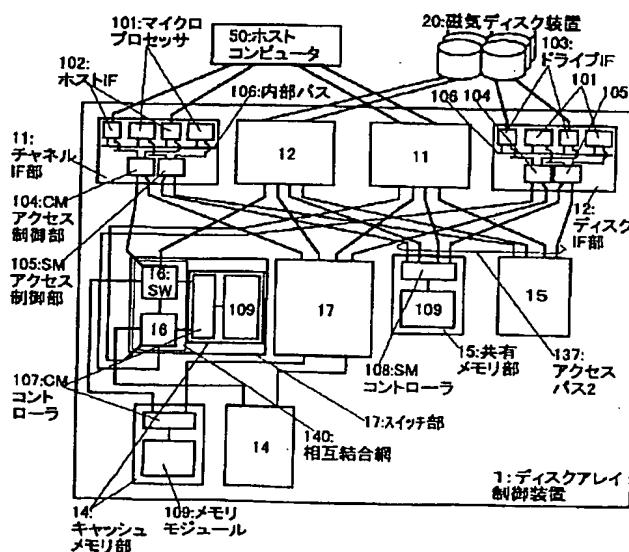
【図9】

図9



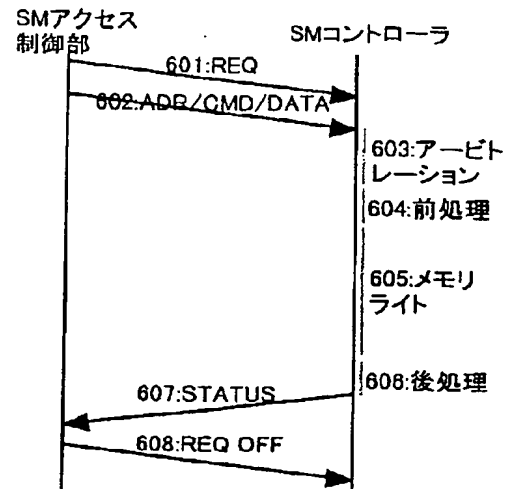
【図11】

図11



【図10】

図10



【図12】

図12

